



Enquête descriptive

Dr BOUNTOGO

Mars 2017

Objectif

- Connaitre les principes des méthodes de sondage
- Savoir calculer le nombre de sujets nécessaires dans une étude épidémiologique descriptive transversale
- Savoir généraliser les paramètres de l'échantillon à la population
- Savoir interpréter les résultats d'une étude écologique ou corrélationnelle

Méthode de sondage

- Problème: Trouver la prévalence de HTA dans la ville de Ouagadougou
 - Difficulté logistique d'avoir accès à toute la population
 - Difficulté financière de contacter toute la population
- Solutions
 - Prendre un échantillon de la population
 - Inférer les résultats à toutes la population
- Comment
 - Echantillon représentatifs (aléatoire)
 - Bonne précision (taille)

Définition de concepts

- Unité statistique
- Population
- Population source
- Population cible
- Inférence
- Sondage ou échantillonnage
- Aléatoire

Comment faire un sondage

- Définir la population à laquelle on s'intéresse: unité individuelle (enfants – 5ans) collective (salle de classe)
- Liste des unités ou comment accéder à ces unités

Sondage aléatoire simple

- Tirage au sort des unités statistiques
- Liste des unités
 - Exhaustive
 - Sans doublon
 - Contact des unités
- Avantages
 - Simple
 - Probabilité individuelle d'être dans l'échantillon = $f = n/N$
 - analyse statistique classique
- Inconvénients
 - Difficulté d'avoir une liste exhaustive

Sondage stratifié

- **Principe:** La base de sondage est divisée en groupes homogènes (strates) selon un critère lié à la variable à estimer
- **Intérêt:** vise à réduire la variance des estimateurs, en améliorant la répartition de l'échantillon entre les différentes strates
- **Avantage:** avoir des individus de chaque strate, sur-représenter des strates plus petites

Sondage en grappes

- **Principe :** Le sondage en grappe consiste à tirer au sort non pas directement un individu, mais des unités collectives (on prend tous les individus de la grappe)
- **Avantage:** réduit le coût en limitant le nombre de visites
- NB: si on est obligé de sélectionner des unités à l'intérieur des unités collectives on parle d'échantillonnage à 2 degrés, 3 degrés.....

Sondage systématique

- **Principe:** pas de base de sondage mais on dispose des information sur la cadences des unités statistiques (malade dans une unité d'urgence)
- **Pas de sondage:** calculer à partir de la taille de l'échantillon et la taille de la population $p=n/N$
- **Choix aléatoire** du premier sujet à tirer et les autre sujet sont tirer en ajoutant chaque fois le pas.

Méthode des quotas

- **Principe** : pseudo-aléatoire
- Attribuer un nombre de représentant par groupe.
- **Avantage**: facile, représentation des minorités
- **Limite**: difficulté d'inférence

Comment tirer au sort

- **Principe:** savoir n et N
- **Avoir** une table des nombres aléatoires
- **Choisir** la direction du tirage
- **Tirage**
- **Autre:** logiciel Excel, stata, SPSS, Epi-Infos

Calcul des paramètres

- **Paramètres sondage aléatoire simple**

- **Pour une variable quantitative**

- $m = \frac{\sum_{i=1}^n x_i}{n}$ avec m = moyenne d'échantillon n = taille de l'échantillon x_i = valeur de la variable pour l'individu x_i

- $s^2 = \frac{\sum_{i=1}^n x_i^2 - \left(\frac{\sum x_i}{n}\right)^2}{n-1}$ variance de l'échantillon

- **Pour une variable qualitative à 2 modalités**

- Proportion $p = n/N$ avec $q = 1-p$
 - Variances $2 = pq/n$

- **Sondage stratifié**

Calcul des paramètres

- **Sondage stratifié**

- Moyenne $\mathbf{m}_s = \frac{\sum N_i m_i}{n}$ avec $N_i =$ effectif de la strate i dans la population et m_i moyenne estimée dans la strate i $n =$ taille total de l'échantillon

- Variance $\mathbf{m}_s = \sum \frac{N_i^2}{N^2} (1 - f_i) \frac{s_i^2}{n_i}$ avec $N =$ taille de la population s_i^2 variance dans la strate i et n_i effectif de la strate i dans l'échantillon

- **Pour une variable qualitative**

- **Proportion** $\mathbf{p}_s = \frac{\sum N_i P_i}{N}$ avec $P_s =$ proportion estimée sur l'échantillon N_i taille de la strate i p_i proportion dans la strate i et N taille de la population

- Variance $\mathbf{p}_s = \sum \frac{N_i^2}{N^2} (1 - f_i) \frac{p_i q_i}{n_i}$ avec n_i effectif de la strate i dans l'échantillon

Résumé

Sondage aléatoire simple	Sondage en grappe	Sondage stratifié
Base de sondage = liste complète des sujets	Base de sondage = liste des grappes	Liste complète des sujets dans les strates
Numéroter le sujet	Numéroter les grappes	Numéroter les sujet par grappe
Statistique simple	Pratique	Gain de précision
Suppose d'avoir la liste	Analyse statistique complexe	Analyse statistique plus complexe que dans aléatoire simple

Interprétation des résultats

- **Taux de participation et non réponses**
 - Pendant l'étude: pré-test
 - Lors de l'exploitation des résultats: on-répondants par rapport aux répondants
- **Interprétation des résultats**
 - Paramètre de position et un de dispersion (moyenne et ET)
 - Echantillon doit être REPRESENTATIF de la population qu'il est censé décrire (individu ont même probabilité d'être inclus)

Nombre de sujets nécessaire dans une enquête descriptive transversale

- **Justification de l'échantillonnage (sans objet)**
- **Précision de la mesure**
 - L'erreur (ou le biais) de mesure, a relire a la validité de cette mesure
 - Le biais de sélection de l'échantillon
 - L'erreur aléatoire
- **Fluctuations d'échantillonnage**
 - C'est une variation du paramètre d'un échantillon à un autre
 - Donc la seule valeur du paramètre ne saurait être la vraie valeur
- **Intervalle de fluctuation**
 - $U_A \pm Z_{\alpha/2} \sqrt{\sigma_A^2 / n}$ U_A la moyenne de la variable A (quantitatif) et σ_A ET et **Condition d'application** normalité
 - $P_A \pm Z_{\alpha/2} \sqrt{P_A Q_A / n}$ P_A la proportion de la variable qualitatif A ($Q_A = 1 - P_A$) et **Condition d'application** : $n P_A \geq 5$ et $n Q_A \geq 5$
 - **NB** P_A et U_A sont des paramètres de la population donc non connus

Nombre de sujets nécessaire dans une enquête descriptive transversale

- **Intervalle de confiance**

- $m_A = U_A \pm Z_{\alpha/2} \sqrt{\sigma_A^2 / n}$ $\rightarrow U_A = m_A \pm Z_{\alpha/2} \sqrt{\sigma_A^2 / n}$ or on ne connaît σ_A^2 mais peut être estimé par $\frac{N-n}{N} S_A^2$ or cette valeur ≈ 1

- Donc $\rightarrow U_A = m_A \pm Z_{\alpha/2} \sqrt{S_A^2 / n}$

- **Éléments du calcul du nombre de sujets nécessaire**

- **Facteurs liés à la taille de l'échantillon**

- Variance
- La valeur attendue du paramètre mesure : déterminée par l'investigateur
- précision de la mesure

- **Calcul la taille d'un échantillon aléatoire simple**

- Variable quantitative

- $n = Z_{\alpha/2}^2 \frac{\sigma^2}{\xi^2}$
- $n = Z_{\alpha/2}^2 \frac{Pq}{\xi^2}$

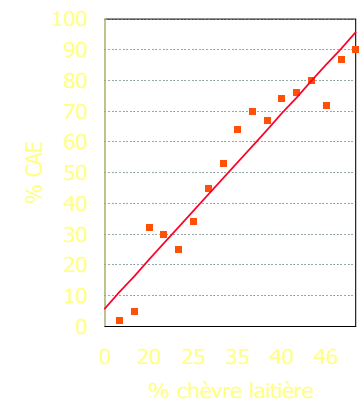
Étude écologique

- Exploratoire
- Comparaison de groupes
- Tendence temporelle
- Comparaison temporelle et de groupes



Étude écologique

- Issue (Taux d'incidence)
 - Numérateur
 - cas incidents si possible
 - définition constante
 - Dénominateur
 - même population que numérateur
 - population à risque
- Exposition
 - supposée égale pour tous
- Association
 - Régression



Merci